FDA Science Board Subcommittee: Review of the FDA/CDER Pharmacovigilance Program

Prepared for the FDA Science Board
May 2011

Contents

FDA Science Board Subcommittee

Review of the FDA/CDER Pharmacovigilance Program

I. Subcommittee Membership

Chair:

Stephen P. Spielberg, MD, PhD
Marion Merrell Dow Chair in Pediatric Pharmacogenomics
Director, Center for Personalized Medicine and Therapeutic Innovation
Children's Mercy Hospital
Kansas City, MO
Member, FDA Science Board

Members of Committee:

Russ B. Altman, MD, PhD
Chair, Department of Biomedical Engineering
Stanford University
Palo Alto, CA
Member, Science Board

Garrett FitzGerald, MD
Professor of Medicine, and Professor and Chair of Pharmacology
University of Pennsylvania School of Medicine
Philadelphia, PA
Member, Science Board

Andrew Bate, PhD
Senior Director, Epidemiology
Pfizer, Inc.
New York, NY

David Madigan, PhD
Professor of Statistics
Columbia University
New York, NY

June M. Raine, MD
Director of Pharmacovigilance and Risk Management
Medicines and Healthcare Products Regulatory Agency
London, UK

Dan Roden, MD
Professor of Medicine and Pharmacology
Assistant Vice-Chancellor for Personalized Medicine
Vanderbilt University School of Medicine
Nashville, TN

Paul Watkins, MD
Director, Hamner-University of North Carolina Institute for Drug Safety Sciences
Research Triangle Park, NC

II. Initial Charge to the Science Board

Charge to the CDER Pharmacovigilance Review Subcommittee

Post-market drug adverse event (AE) surveillance is a core FDA/CDER pharmacovigilance (PV) function supporting the agency's efforts to ensure the safety of drugs and therapeutic biologics. CDER's pharmacovigilance program uses several processes, methodologies and data sources to detect, characterize and prioritize serious adverse events due to drugs and therapeutic biologics.

Charge to the FDA Science Board: The FDA Science Board is charged with conducting a review of CDER's current and planned PV practices. Review objectives include:

- Review the current sources of AEs as well as the processes and analysis tools FDA/CDER uses for identifying safety signals.
  - Regarding the data sources utilized for surveillance, what should FDA/CDER do to improve the collection of high quality spontaneous AEs?
  - Are there changes needed in how FDA/CDER identifies and evaluates significant AE signals in a timely and consistent manner?
    - How should FDA/CDER validate data mining methods and tools?
    - What should FDA/CDER do to improve and encourage the use of analytic tools such as data mining?
  - Is there specific expertise needed in FDA/CDER to support improvement of existing surveillance strategies or development of new strategies?

- Review the current and potential use of specialized cohorts to detect and evaluate AEs and provide input on how they might be incorporated into the FDA/CDER surveillance system.

- Review the current use of public-private partnerships and partnerships with other Federal agencies in the area of PV.

    - Provide input on the potential benefits of establishing collaborations with groups such as the NIH, CDC and private stakeholder clinical trials networks like the AIDS Clinical Trials Network to improve AE collection and quality.

    - What other types of partnerships should be considered for PV?

- Identify emerging science areas most critical for enhancing the capabilities of FDA/CDER pharmacovigilance programs to detect AE's in a range of datasets.

- Identify opportunities to engage multiple external partners for the purposes of encouraging PV research, improving and standardizing PV practice and training, and harmonizing PV policies and practice guidelines between private industry, international collaborators and other PV stakeholders.

As we initiated the review, the complexities of detecting, analyzing, and acting upon adverse event signals became ever more apparent. The Subcommittee envisioned a process leading to recommendations to enhance proactive pharmacovigilance science at FDA, focusing heavily on the spontaneous reporting system. We also recognized that analysis of spontaneous reports, while a current and likely future major source of information on adverse drug effects, is but one aspect of the complex and evolving science of drug safety. It is apparent that FDA needs to develop a systematic approach to scientific validation of a suite of tools to identify, confirm, and act upon potential AE signals in a timely manner. This includes an informed interpretation of preclinical and clinical data focused on potential mechanisms of AEs that needs to be integrated with refinements of more traditional approaches to pharmacovigilance. In evaluating potential AE signals, all independent lines of evidence – mechanistic understanding from pre-clinical and clinical studies, human pharmacogenetics, randomized clinical trials, observational and epidemiologic studies, need to be sought. There is a crucial need to define the benefits, limitations, and biases of each approach, applied to different kinds of situations and outcomes. As an example, different approaches may be needed to detect and analyze signals for "common" versus "rare" outcomes. As shown in the figure below, there are major differences in approach needed for a drug associated with an increase in a common adverse common event (e.g., heart attack or stroke) vs. detecting a rare and unusual outcome (e.g., Hepatosplenic T-Cell Lymphoma or HSTCL associated with certain immunosuppressive therapies). Similarly, different approaches may be needed for situations where there is a strong mechanistic rationale to expect certain outcomes, vs. evaluating an outcome where "biologic plausibility" is much less certain.

What is FDA procedure for discovery and confirmation of AEs?

	Rare AE with important clinical consequences	Common AE not picked up in Phase III that is clinically important
Mode of discovery	?	?
Mode of followup and confirmation	?	?

Assumption: methods to detect rare signals and common signals may not be the same. "Rare" and "Common" cutoff values are debatable, but clearly different procedures and tests apply to AEs that are 1/10,000 vs. 1/100.

Scientific insights drawn from basic pharmacology/toxicology, pre-clinical studies, mechanistic studies in humans, controlled clinical trials, observational studies, analysis of spontaneous reports, and active surveillance (e.g., Sentinel) all should inform establishing and updating best practices for efficient detection of and action upon AE signals. No one approach will adequately address the public health needs of assuring safe/effective therapeutics, and there remains a major need to apply scientifically rigorous methods to maximize FDA's effectiveness in safety assessment and regulation. While not the direct purview of our review, we also recognize that conversion of safety information into labeling that improves physicians' ability to recognize and understand the risk of possible drug-induced symptoms, and to respond appropriately, requires consistent, well-documented description of AEs as well as quantitative risk assessment. This is a major rationale for assuring that spontaneous reports and other sources of input into the AE system have such clinical precision to support safe and effective use of medications in the real world.

Members of the Subcommittee brought expertise in many of these arenas and a perspective for recognizing the need for comprehensive, integrative approaches to the science of drug safety.

III. Review Process

At the outset, the Subcommittee was provided with a detailed document from the Office of Surveillance and Epidemiology (OSE) in CDER. This document outlined the regulatory responsibilities and workload of OSE, the organizational structure and expertise within OSE, an overview of the adverse reaction reporting system (AERS), an overview of pharmacovigilance practice in OSE, and an assessment of data mining analyses of AE reports coming into the AERS data base. Following initial review by the committee, we met weekly by telephone conference, reviewing each aspect of the overview with OSE leadership and staff. The subcommittee then did a full day site visit (October 26, 2010) at FDA, meeting with senior OSE leadership to understand the strengths, weakness, and challenges faced by OSE. We met with several groups of safety evaluators to understand their daily activities/responsibilities, and with Dr. Janet Woodcock (CDER), Dr. Darrell Abernethy (Office of Clinical Pharmacology, Associate Director for Drug Safety) to discuss integrated, science/mechanism based approaches to drug safety), and Dr. Lisa Mathis (Associate Director, Pediatric and Maternal Health Staff, to discuss subpopulations). We had a chance to see current IT tools, and to discuss time commitments and mandated regulatory activities. We heard about visions to enhance pro-active pharmacovigilance, and to improve FDA's ability to efficiently detect, analyze, and act upon AE signals.

We then continued with phone conferences, focusing on implementation of the new FAERS system and its relationship to data mining, reviewing technical and scientific aspects of its implementation with FDA staff. We had the opportunity to review multiple documents and publications by OSE staff (including work outlines for design and implementation of FAERS (FDA Adverse Event Reporting System, FAERS, Systems Requirement Specifications, Nov. 3, 2010), and initial attempts to define new FDA approaches to data mining), as well as a GAO report of post-market surveillance activities at FDA (GAO-10-68, Nov. 2009), and postings from an ongoing IOM review of drug safety. We also reviewed and discussed aspects of the Observation Medical Outcomes Partnership (OMOP) (including presentations of preliminary data from a conference on January 11, 2011), Sentinel, and specialized AE reporting networks.

IV. Sources of Data on Adverse Drug Effects

- **Spontaneous reports**

The number of spontaneous AE reports submitted to FDA continues to grow, with over 600,000 reports in 2010, heading for over 700,000 in 2011. There was considerable discussion by leadership and staff of OSE (as well as by members of the Subcommittee) that there are major problems with the quality of spontaneous reports submitted to FDA (either direct reports, or as the case for the majority of reports, submitted to drug sponsors and subsequently transmitted to FDA). Government estimates of time required to fill in FDA Form 3500 describing an adverse event (AE) suggested an average time of 36

minutes, a very long time indeed for busy practitioners, be they physicians, nurses, or pharmacists. Most reports lack vital information on the "phenotype" of the AE (clinical findings and laboratory findings), drug and doses, confounders, time course, outcomes, and even basic demographics of the patient. Much of this information is present in electronic health records, and can be improved by focused effort to assure critical information is collected, recorded and submitted. We reviewed AE reports of a possible hepatic event in the spontaneous report system and similar reports submitted through the DILIN (The Drug-Induced Liver Injury Network) (see below), and the contrast was stark, the latter containing richness of information characterizing the AE and allowing more accurate description of a specific drug-related process, while most spontaneous reports lack such detail. Analysis of AE reports (see below) by various statistical techniques can lead to invalid conclusions based on inconsistency and inadequacy of data provided in most reports. Similarly, even evaluating the number of submitted reports of a specific AE can be biased by publicity and litigation. Clinical precision of reports is vital in providing regulators a medically valid basis for labeling that can be effectively used by practitioners in the real world.

- **Specialized cohorts, networks**

There currently are a number of networks established nationally and internationally focused on ascertainment and evaluation of adverse drug effects. Some, such as the International Severe Adverse Events Consortium (iSAEC) and the Drug Induced Liver Injury Network (DILIN) focus on specific organ targets or patterns of adverse drug reactions. They have served as major sources of high quality AE reports as well as collaborative research efforts into the mechanisms, predisposition (pharmacogenomic and other), and pathogenesis of adverse drug reactions. Other networks focus on therapy of specific disease states, such as the Children's Oncology Group (COG). Finally, there is a need for additional networks focused on special patient populations – e.g., children (for example the Canadian Pharmacogenomics Network for Drug Safety (CPNDS)), women, and the elderly. It is clear that AEs in developing children are often quantitatively and qualitatively different than in adults, and risk factors such as genetic predisposition are complicated by the very processes of growth and development. Discussions with the Office of New Drugs (OND) Pediatric and Maternal Health Staff suggest a need for targeted surveillance of pediatric and other special populations, as well as a need for internal expertise at FDA to assess special population signals.

The DILIN system is an example of a high performing network. The network consists of 10 large academic medical centers across the country and is funded by the National Institutes of Diabetes, and Digestive and Kidney Diseases. DILIN has now enrolled over 900 people who have experienced drug-induced liver injury. Extensive clinical data are obtained on each subject and adjudication of cases is performed by at least three hepatologists experienced with drug-induced liver injury. Tissue samples including genomic DNA, serum, urine and liver (when available) are obtained from each subject and made available to researchers. Each case is reported through the AERS voluntary

reporting system once adjudication is completed. Genetic studies are underway and have revealed unsuspected mechanisms that should lead to tests that can assistant in the diagnosis drug-induced liver injury. Data generated can help improve characterization of a specific drug-related adverse event phenotype, and thus help inform how best and most comprehensively to gather the right kinds of data from spontaneous reporting or from active drug surveillance. This is important both for analysis of AE signals within FDA, and for providing practitioners accurate information through labeling to be able to identify possible adverse reactions in their patients as rapidly as possible, alter dose or stop therapy, and possibly prevent further drug-induced disease. Networks such as DILIN also do mechanistic studies that can lead to biomarkers to improve diagnosis, prediction, and prevention of hepatic injury risk. Staff from CDER serve as ad hoc members of the Steering Committee of this network.

- **Controlled clinical trials, observational studies, post-market commitments**

Under the FDA Amendments Act of 2007 (FDAAA), new section 505(o) of the Federal Food, Drug, and Cosmetics Act authorized FDA to require certain safety-related postmarketing "studies and clinical trials". This area was not within the scope of the Subcommittee's review. Members of the Subcommittee, however, did comment on the need for standardized definitions and clinical work-up of suspected adverse outcomes of pharmacotherapy. The need for increased precision of clinical and laboratory description of patients in spontaneous reports, by specialized networks, in active surveillance programs, and in these types of studies is equally important**.** Consistency of evaluation of AE signals and collection of data across clinical trials through post-market surveillance will facilitate accurate post-market reporting, and, should a signal be detected, permit returning to clinical trial data to see if AEs detected post-market can be found in the clinical trial databases. FDA should be pro-active in defining and evaluating safety data standards and content for different types of AEs.

- **Active Surveillance: OMOP/Sentinel**

An appealing vision for the future detection of drug responses, and other healthcare outcomes, is advanced electronic medical (health) record (EMR or EHR) systems that can aggregate and ultimately even start to interpret data obtained from large numbers of patients. Pursuant to the FDA Amendments Act (FDAAA), the FDA is directing significant resources towards the construction of an active adverse event surveillance system, the Sentinel system (http://www.fda.gov/downloads/Safety/FDAsSentinelInitiative/UCM233360.pdf). Specifically, FDAAA instructs FDA to link and analyze data covering at least 100,000,000 patients for the purposes of drug safety evaluation by 2012. The Subcommittee was struck by the unevenness of knowledge about Sentinel among the OSE personnel with whom we spoke. Some expressed opinions that Sentinel would primarily be used for "one-off" pharmacoepidemiology studies, while others believed it

would be crucial for risk identification. Some of this may be due to differences in understanding of time lines for implementation, short and long-term goals. We urge the FDA to develop a clearly articulated vision for Sentinel providing specific goals and a detailed timeline. FDAAA calls for an "active surveillance" system yet no clear definition of the term "active surveillance" currently exists. Of considerable concern to the Subcommittee, is the apparent variability in understanding of how Sentinel and the spontaneous reporting system will interact, and we were not presented with a strategic plan nor concept of how the spontaneous report system will be impacted by Sentinel.

In preparation for the design and implementation of Sentinel, FDA participated in a number of initiatives, including a pilot study to determine if the systematic application of a variety of statistical and epidemiological methods to current observational data could detect pairs of drugs-adverse events that have been validated by other methods. OMOP (Observational Medical Outcomes Partnership) was funded as a public-private partnership under the auspices of the Foundation for the NIH, and included expertise drawn from FDA and other federal agencies, industry, non-profit organizations, and the academic community. The Health Outcomes of Interest study undertaken by OMOP evaluated 53 drug-outcome pairs across 10 large data sources. Preliminary results of the study were presented January 11, 2011 and all data are posted on the OMOP website (http://omop.fnih.org). There are an enormous number of important insights gained from the study – from variability in condition-prevalence (and disease definitions and nosology) and drug use among different study sites, complexity of data analysis, need for validation across all data sets. Two summary tables below highlight the complexities discovered by the exercise:

OBSERVATIONAL
MEDICAL
OUTCOMES
PARTNERSHIP

• An active surveillance system can successfully complement current practice by providing useful evidence to support a comprehensive safety assessment
• No one clear 'best ' method, as it depends on tolerance for false positives vs. false negatives
• Systematic pharmacoepidemiology can achieve:
  – At 50% sensitivity, false positive rate ranges 16%-30%
  - At 10% false positive rate, sensitivity ranges 9%-33%
• Need to be cautious in interpreting results from single method in single database
  – Replication does not necessarily provide complete confidence
• You need a relative risk > 2 to have confidence in result
….detecting effects smaller than 2 will incur higher risk of false positives

OBSERVATIONAL
MEDICAL
OUTCOMES
PARTNERSHIP

• Method performance can vary by data source, drug, and outcome
• Method estimates are sensitive to outcome definitions and parameter settings
• Need to be cautious in interpreting results from single method in single database
  – Replication does not necessarily provide complete confidence
• Need to develop strategies for principled parameter selection and implement comprehensive sensitivity analyses for evaluating the robustness of any findings across:
  – Data source and target populations
  – Method and parameter settings
  – Outcome and exposure definitions
• Additional research across a broader array of test cases is needed to fully characterize expected method behavior to improve confidence in the results that are obtained

As is the case for spontaneous reports, data input quality is vital to interpretation of any results. Accuracy of phenotypic description of outcomes, consistency of work-up and evaluation of patients with possible drug-induced disease, and consistent use of terminology all are key factors. Analytical techniques, their strengths and limitations, are discussed extensively by the OMOP investigators, and point to the need for a variety of different kinds of ascertainment and analysis tools to detect and validate AE signals. Trade offs, for example between sensitivity and false positive signals need to be understood, and relative performance of different types of analyses compared. It is critical that further scientific experiments are conducted to advance understanding of what kinds of techniques are most likely to detect/validate specific types of AEs (by frequency, complexity of outcomes, demographics [e.g., age, gender]). FDA, together with external expertise, must take the lead in explicitly defining input and evaluative quality needed to optimize rapid detection and analysis of data.

The results from OMOP should inform the Sentinel initiative as it evolves. Ongoing research through OMOP, and other targeted research initiatives are needed to validate approaches to active surveillance, determine the robustness of the process for signal detection and evaluation, and form the basis for FDA integration of active surveillance as part of the larger 'tool box' for the detection and analysis adverse drug reactions.

- **Novel sources of data – internet search engines and social networks**

Growing access to electronic connectivity, internet search engines and social networks are having a profound impact on all aspects of society. The use of internet data resources

and social networking raises the potential for internet-based public health signal detection. Searches for symptoms have been suggested as a means for monitoring influenza outbreaks, for example. For example, available data indicate that patients with flu symptoms may start searching the internet for remedies, symptomatologies, and other information well before traditional monitoring mechanisms detect an outbreak. The logs of the major search engines (Google, Yahoo, Bing, etc…) are a potential goldmine of direct-from-consumer health signals. Although the best methods for mining these data are just now being defined, the FDA has a public health interest in this data, and should consider partnering with the companies that provide these search engines to test their utility in drug safety surveillance. The incredible volume of searches by patients may yield important clues that at least can support or refute internal FDA hypotheses about emerging signals. There are important privacy implications of gaining access to these data, but it is very clear that the companies are already maintaining this information and the FDA may have a legitimate public health mandate to work with the companies to create methods for secure and confidential access for the purposes of public health. Search logs may provide early indications of unexpected adverse events, as patients go to these engines wondering whether their symptoms are related to recently added or changed medication lists, and thus literally linking the drugs with their adverse effects via their search terms, thus potentially creating a new type of adverse event reports.

In addition to search engines, social network sites such as Facebook and Twitter offer another example of emerging internet phenomena that may aid the FDA in its drug safety mission. These resources are new and expanding at unprecedented rates, with tens of millions of individuals contributing information. To be sure, these data sources are novel, noisy and methods for their analysis for serious drug-related purposes must be defined. However, they offer an unprecedented opportunity to get electronic access to the population that the FDA serves. As the "early adopters" of these social networking technologies (often teenagers and college students) age, enter the workforce, and increasingly engage the healthcare system, the FDA should have a plan for using these resources to assist in their mission.

Finally, while not fully analogous, aspects of FDA signal detection share some similarities to the signal detection methods of internet-based companies involved in marketing, sales, user preferences, and search refinement and personalization. For example, the Netflix movie rental company recently sponsored a $1M "Netflix Prize" for algorithms that could predict which movies a customer would enjoy. Google researchers are constantly refining methods for ranking the hits on a Google search. Credit card companies have computer scientists creating algorithms to detect credit card fraud. There is little question but that these evolving technologies have the potential to impact every area of our society. Their growth is an inevitable fact of life. In the very least, these new technologies are likely to impact how we recognize, record, transmit, and assess medical data, as well as how we practice medicine. How to harness all of this in the interest of the public health and how to validate signals detected through these approaches will be a huge challenge, but one that must be undertaken. The FDA should consider convening as broad an array of methodologists from statistics, computer science, and epidemiology to address emerging signal detection approaches.

- **Mechanistic pharmacology/toxicology**

There have been major advances in pharmacologic and toxicologic science that can have a direct impact on how the safety of medicines is evaluated. There is no question but that insight into basic mechanisms of action and potential adverse effects can inform and prepare the clinical trials and post-market processes, and there is a major need to assure that such scientific thinking is incorporated into "priming" epidemiologic systems to search for and evaluate potential drug toxicity. Among elements of Translational Medicine and Therapeutics (TMAT) that might be integrated to help predict risk are: pre-clinical pharmacology and toxicology (potential signals, mechanisms, how signals might translate into human risk assessment); pharmacokinetics (PK)and modeling (and modifications of PK by variables including age, gender, diet, disease state, other medicines, pharmacogenetic variants in the human population); pharmacodynamics (target specificity, potential "spill-over" including influence of genetic variants of intended and alternate targets); mechanistic human toxicology and development of biomarkers (e.g., pharmacogenomic, immunologic, other biomarkers, with an ultimate focus on individualized therapeutic and adverse effect prediction). As such, there is a need for continuous and ongoing interaction internally among scientists and regulators within FDA. Much of the progress and capacity for these diverse approaches do and will continue to reside in the academic sector and the NIH, and there is thus the need for structures to integrate the interests of FDA with academic and government research partners. Similarly, there are opportunities to engage across FDA, academia, and the pharmaceutical industry, creating pre-competitive public-private consortia to develop and validate new science (e.g., the FNIH Biomarkers Consortium).

The establishment of a group in the Office of Clinical Pharmacology led by Dr. Darrell Abernethy focused on pharmacologic mechanism based safety prediction should be a positive development. The group is currently small and new in its function, but provides the opportunity to bring together scientists and scientific disciplines to improve the quality of safety assessment from pre-clinical through post-market. The potential of this initiative could be greatly enhanced by creative partnerships with the academic community focused on drug safety, such as is supported by the Medical Research Council at the University of Liverpool, UK. Development of IT structure within FDA to assess and validate different approaches to safety retrospectively and prospectively (including initiatives such as the JANUS clinical data repository) is needed to deal with the large volumes of data generated. Similarly, as international networks are developed to characterize adverse drug reactions and to study their pathogenesis, there should be efforts to link international centers and investigators. One activity already underway is to develop a taxonomy of molecular toxicology targets mapped against drugs.

Many of the TMAT elements, including pharmacogenomics, are rapidly evolving. Their validation and incorporation into safety assessment (and all of FDA's regulatory activities) is complex and will require interdisciplinary thinking and collaboration. The rapidity of development of new knowledge that will change the process of drug development and evaluation (e.g., a focus on mechanistic pathways rather than traditional

diagnoses) is a central challenge to the future of regulatory science. In the realm of safety assessment, it is crucial for that creative, collaborative interactions be fostered among OSE, the Office of New Drugs (OND), the safety group in the Office of Clinical Pharmacology, and the Office of Testing and Research's new Division of Drug Safety Research directed by Tom Colatsky. As well, there is need for established procedures and pathways for evaluation of potential mechanisms of toxicity within FDA, and together with NIH and the academic community. The latter interactions would benefit from establishment of ***Centers for Drug Safety Science***, perhaps as part of NIH efforts in translational medicine and analogous to the MRC center in Liverpool, specifically focused on issues of adverse drug reaction mechanisms, and constituted so as to be able to rapidly partner with scientists within FDA to address emerging issues of drug safety science. Such an initiative is also congruent with the suggestions of the subcommittee of the Science Board (http://www.fda.gov/ohrms/dockets/ac/07/briefing/2007-4329b_02_01_FDA%20Report%20on%20Science%20and%20Technology.pdf) that previously reported more generally on FDA science and which recommended collaborative centers in academia focused on areas of science of relevance to regulation, such as drug safety. With expanding knowledge and sophistication in all spheres from clinical ascertainment, epidemiology, through genomics, the challenge will remain to assure integrative approaches, bringing together all realms of safety research. Such initiatives would resonate with a likely focus of the new National Center for Advancing Translational Science (NCATS) on drug de-risking and repurposing, taking advantage of the NIH peer review process, and the NIH/FDA roundtable which will be located within NCATS. These initiatives provide timely opportunities for consideration of how optimally to organize interdisciplinary drug safety investigation.

It is clear that while FDA does not have the research resources to undertake such studies alone, the agency is in a unique position to set research agendas to support crucial public health needs in drug safety, and to act as a convener to drive research activities. FDA similarly can play a major role in research by creating a series of "gold-standard" data sets against which new technologies and science can be evaluated. Partnerships with the NIH and the private sector, and public-private partnerships, both national and international, will be needed to advance the safety sciences to meet regulatory needs. There is a clear need for partnerships with the NIH to extend beyond drug-discovery and translational science to include drug safety sciences.

V. Data Management:

- **Current systems and procedures in the Office of Surveillance and Epidemiology**

OSE has two divisions of pharmacovigilance, currently with 14 medical officers and 51 safety evaluators. The workload for the Office has increased dramatically over the last years. The number of individual case safety reports (ICSRs), predominantly adverse event (AE) reports but also including medication errors reports, has increased from approximately 250,000 in 1996 to over 600,000 in 2010. The percent of serious AE reports entered into the AERS (Adverse Event Reporting System) has remained relatively

stable, averaging 77% between 2000 and 2009, and percent of deaths as a percentage of serious reports again has been stable at approximately 16% during this time frame. The vast majority of the reports come through drug manufacturers, perhaps 5% or less of reports coming directly to FDA.

As pointed out in a GAO report in 2009 (GAO-10-68), OND and OSE both have responsibilities related to monitoring the arrival of, and evaluating these reports. Manufacturers submit ICSRs both as 15-day reports (of serious and unlabeled adverse events) and as Periodic reports (of serious labeled and nonserious adverse events) that are received quarterly or yearly. The ICSRs are entered into the AERS database upon receipt. AERS queues the 15-day reports to OSE safety evaluators for review once they have been entered into AERS. In addition, OND medical officers receive a weekly listing of 15-day reports that have been entered into AERS. The Periodic ICSRs arrive at FDA with an accompanying descriptive summary of the safety data received by the manufacturer during an inclusive quarter or year. OND medical officers conduct the primary review of this Periodic report summary, while OSE safety evaluators receive a copy. The small percentage of direct-to-FDA reports are also queued by AERS to OSE safety evaluators. The reports in AERS, including 15-day reports, periodic reports and direct-to-FDA reports, are also monitored by means of AERS searches, standard report output and data mining activities, although the latter are often carried out on an *ad hoc* and non-systematic basis (see below).

As pointed out by the GAO report, and discussed by many of the staff with whom we spoke in this review, there often are differences in underlying philosophy, reliance on controlled clinical trials (OND) vs. observational data/epidemiologic approaches (OSE), between the two offices. There are, as well, issues of "ownership" of data and decision making processes. Diversity of approaches to understanding possible drug-induced adverse effects, and open dialogue in the context of well established systems, quality and performance can advance understanding and ultimate decisions. Inconsistencies and poor communication, as well as lack of clear accountability and processes, however, can weaken the review process and lead to delays in action. Section 901 of FDAAA includes provisions for risk evaluation and mitigation strategies (REMS) that require decisions about REMS to be made "in consultation with the office responsible for reviewing the drug and the office responsible for post-approval safety with respect to the drug," which CDER has interpreted as OND and OSE. Section 921 mandates OSE bi-weekly reviews of the AERS data base for any new safety signals of potential serious risk, and MAPP 4151.7 describes responsibilities and interaction of OND and OSE in identifying and managing AE signals, including regular joint safety meetings . These statutory and policy efforts focus on trying to enhance collaboration and clarify responsibilities between OND and OSE. From discussion with staff in OSE, some strains remain between OND and OSE; some review divisions such as Division of Cardiovascular and Renal Products were cited as having worked particularly well at assuring maximum collaboration, but this was not uniform.

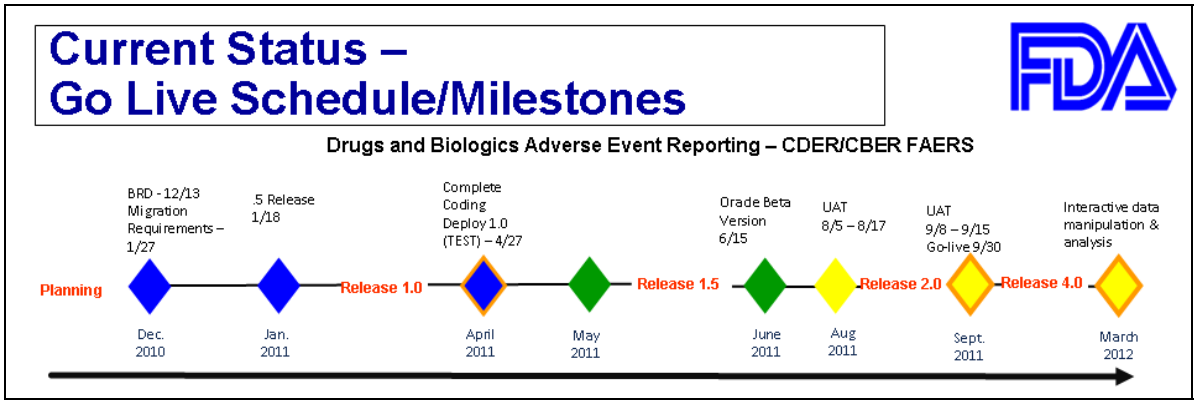
Currently, the workload within OSE is being heavily driven by consultative requests from medical officers in OND, as well as new legislative mandates. The Best Pharmaceuticals

for Children Act (BPCA) specifies 1 year post-exclusivity safety reviews. Staff in the OND Pediatrics and Maternal Health Staff noted the need for more pediatric expertise to advance the scientific base of these reviews, which utilize considerable OSE resources. Similarly, FDAAA mandates 18 month (or 10,000 patient exposures, whichever is later) post-approval safety evaluations for all newly approved drugs. The workload of these mandated reviews and consultations from OND were viewed as substantial and perhaps impeding OSE staff from advancing safety science. It was less clear about the public health value of the mandated reviews. While rational on the basis of what of we know about adverse event detection when compounds go from clinical trials to general population use, it would be of real value to have hard data to assess the utility of these reviews, and to plan personnel and technical needs to support the expanding activities of OSE.

- **Management of AE reports (the AERS data base)**

There was nearly universal agreement that the current AERS data base and IT infra-structure for capturing and evaluating potential AE signals is outdated and inadequate to the tasks of OSE and FDA. The system lacks integrated tools to manage reports, the drug dictionary is duplicative and adds confusion rather than clarity (drug names are often stored in whatever form they are submitted and not mapped to a standard dictionary), and the data base cannot be directly transferred into data analytic tools (including Empirica™ Signal) requiring substantial time and potentially introducing errors moving from one system to another. Some data from Medwatch forms requires manual manipulation prior to analysis, with associated lost time and potential errors. All agreed that the first priority for OSE to improve safety science is to replace the AERS system and to incorporate functionality that will enhance the ability of safety evaluators to do their work. Parenthetically, there is wide-spread variability in how the safety evaluators use AERS in its current configuration, and how, when, and why they undertake data mining (see below), a process which is cumbersome and inadequate with the current IT infra-structure.

The subcommittee review took place in the midst development of the new FAERS (FDA Adverse Event Reporting System) which will replace AERS. A time line for implementation is shown below:



The planned FAERS system will be web-based (vs. client server for AERS), has a dashboard for evaluators to track their multiple products and signals, has alert systems for serious AEs, has an improved dictionary, and ease of duplicate report identification (many reports go both to companies and the FDA). The committee was provided with systems requirements dated Nov. 4, 2010. We were not able, however, to see pre-release versions of FAERS. Crucial to the future development of safety science within OSE, many questions remain about the ability to apply analytic tools such as Empirica™ Signal, and if FAERS will truly meet its anticipated impact on improving IT and workflow within OSE and throughout CDER. Speaking with many safety evaluators, there was general enthusiasm for the FAERS initiative given many of the problems associated with AERS, but concerns remained about implementation. Less certain is a systematic approach to defining the strengths as well as ascertaining the limitations of the system, to establish realistic expectations of what the FAERS can and cannot do, where it can be modified with experience, and how it fits into other safety initiatives with CDER and OSE. The dearth of documentation on FAERS available during the Subcommittee's deliberations precluded our evaluation of FAERS functionality, its strengths and limitations, and of how data mining and other analytical processes will be implemented through the new system. Critical issues of how FAERS and data mining will be used (see below), remain uncertain.

- **Data mining**

As mentioned above, there were inconsistencies among safety evaluators on when, how, and why to do data mining, indeed if the role of data mining is signal detection, signal verification, or simply a means of handling the increasingly large volumes of individual reports coming into the system. Similarly, from a technical point of view, while it was stated that the FAERS system would have much functionality with respect to data mining, the details remain uncertain. Indeed, the design of the IT systems should flow from a scientific assessment of the role of data mining – for example, whether it will be fully operational "in the background", generating reports on a periodic basis for all compounds, or be selectively applied based on some "signal" seen first in the FAERS system. A recent OSE draft document (2/19/2011) looking forward to advancing data mining at FDA included the following conclusions:

1. Establish a data mining/information science center in OTS. Data mining and related information science programs need a home where regulatory research, evaluation and testing may occur before new data mining and related tools are deployed within CDER.

2. Integrate the data mining group's short term goal of systematic use of data mining as part of a larger medical informatics strategy for CDER. Data mining approaches and tools may be used with many databases with basic, clinical and epidemiologic

information. To encourage a seamless interface between the numerous CDER data sources, the data mining staff would identify and employ the best scientific tools and methods to evaluate the various datasets collected throughout CDER.

3. Create and support a multidisciplinary staff of scientists to accomplish the short and long term CDER data mining needs.

4. Establish a scientific forum for data mining and medical information management experts to facilitate scientific interaction between FDA, and other federal data mining experts. A long term goal would be to enhance technical discourse between CDER, other FDA partners, federal collaborators and non federal experts concerning the latest findings for the best data mining and informatics practices and tools.

These conclusions make some sense, and emphasize the need for scientific validation of data mining in many contexts throughout the drug development process and within CDER. There are many caveats about the science of data mining and situation-specific decisions made regarding: specific tools used (and under what circumstances), how analysis of spontaneous databases "fits" with data gathered from other sources (e.g., Sentinel or clinical studies), setting thresholds for action, and how the data will be interpreted and used in a regulatory context. As demonstrated in the OMOP study, small changes in parameters and tools used in analysis can lead to very different results. The strengths and weaknesses of different analytic methods and IT tools needs to be evaluated and made transparent in the decision making process. Similarly, transparent discussion of the regulatory significance of sensitivity, specificity, false positives and negatives need to be discussed in the context of any of the approaches. Looking towards a more integrative approach to using multiple data streams - FAERS, Sentinel, networks, new sources (e.g., social media) - there is a need for a comprehensive approach to data mining/analysis across all these streams. This is a crucial time to re-evaluate data analysis tools. CDER is currently using a data mining algorithm, multi-item gamma Poisson shrinker (MGPS), that dates from the late 1990's. Despite this, operating characteristics and performance on AERS remain uncertain, and there has been extensive new research of other approaches in this area. Validation of how analytical tools perform across data streams is critical to planning a comprehensive approach to drug safety science at this time.

VI. Recommendations

The goal of this subcommittee review was focused on how to advance the "science of pharmacovigilance" within FDA. That goal is one critical aspect of enhancing optimum approaches to timely adverse event detection, analysis, and regulatory action. Everyone with whom we interacted at FDA is committed to this goal, and supported more proactive AE ascertainment and evaluation. It is apparent that science is changing at an ever more rapid pace. Progress in the arena of drug safety will rely on ***an integrative scientific approach*** – from basic and clinical pharmacology/toxicology, genomics, and all science

relevant to the progressive personalization of medicine, controlled clinical trials, observational studies, active surveillance (e.g., Sentinel), spontaneous AE reporting.

A cardinal element of this report is to encourage FDA to combine a refinement of current approaches to pharmacovigilance with the incorporation of independent streams of data relating to mechanisms – studies in cells and model systems, evoked phenotypes and genomic analyses in relatively small numbers of individuals, humans genetic, and randomized trials. Forward progress will be enhanced through establishing and incenting robust interactions with expertise in the academic and private sectors.

IT infra-structure must support interaction across all these arenas. No one type of analysis will always yield "the right answer"; each has strengths, weaknesses, and limitations in detecting and analyzing different kinds of signals, and this needs to be accepted "up front". While no one approach will serve as a "gold standard" in every circumstance, convergence of data from different types of analyses may help improve confidence that a "signal" is truly a safety concern. Experimental approaches to validation such as OMOP will be required to validate methods and elucidation of limitations of any specific approach.

The Subcommittee identified major issues with respect to current AE databases and data mining, uncertainties with respect to implementation of the new FAERS database and its interface with data mining, and a need for thoughtful evaluation and implementation of data analysis platforms across increasingly diverse streams of data input – spontaneous reports, active surveillance, novel sources such as social networks.

With the accelerating pace of science, and the dual needs for FDA to stay at the cutting edge of new science while simultaneously taking regulatory action in real time, we propose the following overall recommendation:

- **General Recommendation**

FDA should establish a standing committee, possibly a subcommittee of the Science Board, dedicated to improving signal detection and safety evaluation. The committee would consist of members with expertise across the spectrum of adverse reaction prediction, AE signal detection, evaluation, validation, as well as members with expertise in mechanistic and clinical pharmacology/toxicology and human therapeutics. The overall goal would be to help FDA develop an integrative approach to safety science, and to develop structures and procedures within FDA to assure successful application of science to regulatory action. Similarly, recognizing that scientific advancement in the field of drug safety will require partnerships with the FDA, academia, and private sectors, the committee can act as a facilitator for such relationships. Specific short-term goals would focus on improving AE report quality, on assisting FDA with optimal utilization of FAERS for AE signal detection, on articulating a vision for Sentinel and how it co-exists with other data sources, and develop integrative, validated analytic tools across the spectrum of data streams that can help identify AE signals. Development of standardized policies and procedures for data analysis based on these considerations is

vital. Similarly, the committee would assure maximum consideration of the strengths, weaknesses and caveats surrounding all methods, and suggest how evolving tools (including rapidly advancing arenas such as social media) can be integrated into a comprehensive approach to timely ascertainment of and action upon safety signals.

- **Long-Term Recommendation**

FDA needs to be proactive in developing, evaluating, and implementing a "suite" of tools to detect, evaluate, and act upon adverse effects of pharmacotherapy. This will be an evolving process as science and technologies advance. There needs to be a systems-wide approach across CDER with clear "ownership" and responsibilities, based on integrated application of science and personnel. **IT support linking different tools (data acquisition, storage, analysis), and different components of CDER is vital.**

FDA does not, and nor should it have all the research capacity within the agency to do such studies. FDA, however, is uniquely positioned to set the research agenda to support its public health missions in drug safety, as well as providing "gold-standard" data sets that can be used for method validation. Successful scientific evaluation of technologies and approaches to AE identification and evaluation across multiple data streams will require that FDA partner with the NIH, academia, and the private sector (pharmaceutical and companies involved in developing EHRs). Implementation of evolving methodologies and coordination among different parts of the agency will be the responsibility of FDA. The above recommended standing committee should assist FDA in the scientific and administrative implementation of advancing drug safety science.

- **Short-, Medium-Term Recommendations**

The first three recommendations speak to the urgency of immediate action needed on FAERS implementation, an integrated approach to data mining, and establishing a clear vision for the interplay of Sentinel with other data streams. Subsequent recommendations emphasize areas of opportunity and need divided among data sources, organizational practices and data management.

  - **Recommendation 1**

FDA needs to take advantage of the opportunity of creation of FAERS to optimize its functionality. We were only able to review the broad outlines of the new FAERS system. Improved drug dictionary, identification of duplicate reports, reviewer dashboards, etc. were mentioned as improvements over AERS, but this is a crucial time to identify limitations up front that may be fixed prior to implementation. One gap identified was uncertainty of how FAERS will interface with data mining systems. This

needs to be clarified up front – will the system routinely undertake data mining on some periodic basis, under what specific circumstances would reviewers want to initiate secondary data mining analyses based on their review of FAERS data, etc. At the very least, external review of pre-introduction versions by outside experts, including the standing committee would be valuable. This will also be important in design new electronic input methods – e.g., smart forms derived from EHRs. External inputs into FAERS optimally should provide a seamless, user friendly port of entry, with clearly specified safety data standards, and IT simplicity in submission into the new system.

- **Recommendation 2**

FDA urgently needs to develop a working group, of internal and external advisors, to define system functionality for data mining with FAERS, as well as other data streams for AE reporting. FAERS needs to be built so as to facilitate use of current data mining tools, and be adaptable as new methods evolve. The performance of data mining activities using FAERS needs to be supported by suitable quality measures. Determining the performance characteristics of data mining is essential to provide a benchmark for comparison in the assessment of emerging data mining algorithms. Ongoing transparent assessment of emerging algorithms is essential to ensure that FA can continue to use the best possible tools in a rapidly evolving field.

The working group needs to define technical specifications and IT requirements, select and evaluate specific data mining tools, and most importantly to establish SOPs for who, why, when, and how data mining is to be done. Again, we heard a range of understanding and comments on data mining. OSE's document on data mining 2/09/2011 represents a beginning towards these goals. It cannot be over-emphasized that spontaneous AE report collection, evaluation, signal detection, and evaluation should not be considered in isolation. Rather the goal is to create a highly functional, integrated system, including multiple tools and staff with various backgrounds, with clear coordination and SOPs. OSE is a logical and appropriate Office to take the lead, but will need strong support from FDA leadership, and a collaborative spirit throughout FDA to advance the science and methodologies of safety assessment and regulation.

- **Recommendation 3**

FDA needs to articulate a clear plan for Sentinel and how it will co-exist with other tools. Current uncertainties and varying conceptions of Sentinel may be hindering progress. This plan should address basic research and development needs to support Sentinel in the coming years, as well as specifying the interactions of Sentinel with the spontaneous adverse reaction reporting system and other data streams for adverse drug effects.

- **Recommendation 4**

FDA should take the lead in the development of "smart form" AE reports, easily adapted from EHRs, and easily transmitted to FAERS. All agree about the need for *high quality* input into the system. Improving EHR capacity to recognize and document AEs is crucial priority for the spontaneous reporting system, for the future of Sentinel, for providing accurate diagnostic information about AEs in drug labels, and for advancing research into adverse reaction mechanisms and regulatory science. Some recent publications suggest potential approaches to the generation of information rich, clinically meaningful, and phenotypically accurate AE reports. Generation of high quality spontaneous reports cannot remain a 36 minute task. Autopopulation of much data directly from the EHR can save time and increase reliability of content. Determination of content of reports for different kinds of AEs should take advantage of expertise and knowledge gained from drug-induced disease networks (e.g., iSAEC, DILIN, ITCH, etc.). Content of reports for special populations (e.g., children) needs to be modified for factors relevant to those populations (e.g., effects of growth and development on drug PK, PD, and effects of drugs on growth and development). FDA should consider partnering with the industry, academic community, and other government agencies such as the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) and Centers for Medicare and Medicaid Services (CMS), and companies providing EHR technology. Making AE detection and communication a priority for accreditation and reimbursement may provide an impetus to more rapid implementation. Similarly, novel approaches should be explored towards enhanced interaction with AE reporters. Positive, timely and relevant individual feedback and recognition as well as novel means of providing enhanced clinical details using IT approaches should be considered.

- **Recommendation 5**

FDA should play an active role in establishing and supporting adverse reaction networks. This includes current networks focused on specific drug-induced adverse outcomes where clinical ascertainment, precise definition of phenotypes, submission of data to FDA and regulatory authorities, and research into mechanisms are integrated activities.

In addition, networks focused on specific demographic populations (children, women, elderly), as well as management of specific disease states (e.g., Children's Oncology Group) should be supported and established to provide additional specialized AE input, and to act as community-based networks to undertake a variety of studies to test hypotheses generated by spontaneous reports or through Sentinel. Special population input will require comparable internal FDA expertise for evaluation and action. FDA should set specific research agendas for such networks based on regulatory science need, and partner with the NIH, AHRQ, industry, and the academic community, creating public-private partnerships to enhance the "input side" of drug safety science.

- **Recommendation 6**

Together with the standing safety committee, FDA needs to address the scientific basis of AE detection and evaluation using a variety of methods (spontaneous reports, Sentinel, targeted studies), and develop a strategy for integration of science from basic mechanisms through epidemiologic investigations. FDA has major initiatives underway with introduction of FAERS, new data mining approaches, Sentinel, etc., but these are viewed very differently across groups with whom we spoke. How each is used and integrated into a comprehensive approach to AEs should be science based (OMOP is a good example of an experimental approach). There is a risk to different approaches developed independently or in a fragmented manner, and it is apparent that each approach will have strengths and weakness in detection, evaluation, and in supporting regulatory action. This integrative process needs to be implemented immediately to avoid creating goals, systems, and tools that do not maximize FDA's role in drug safety.

- **Recommendation 7**

FDA should take a leadership position in integrating scientific approaches directed at elucidating the mechanisms of drug action into detection and prediction of risk, and foster the development of structures to enhance communication of newly emerging science relevant to missions the agency. Scientific understanding of potential mechanisms of adverse drug effects, e.g., the role of human pharmacogenomics, is advancing at a rapid pace. FDA has already demonstrated leadership in incorporation of new knowledge into labeling (e.g., warfarin), and establishing a new safety program in the Office of Clinical Pharmacology. This should be supported and encouraged. Difficult regulatory decisions based on ever more rapidly evolving science will require broad and transparent collaboration and evaluation of the validity of that science in identifying risk and preventing adverse drug effects. While the vast majority of scientific developments will likely come from the NIH, academia, and the private sector, FDA is uniquely positioned to convene expertise to envision how scientific advances are translated into regulatory science and action. If new scientific discoveries are to improve the public health, and the outcomes of patients treated with ever more novel approaches to therapeutics, regulatory science and the FDA must be at the cutting edge, wisely incorporating new knowledge into practical approaches to improving the evaluation of therapeutic interventions.

- **Recommendation 8**

FDA needs to take further action to clarify responsibilities in safety assessment among key internal stakeholders – from NCTR, to OND, to OSE – to assure maximum utilization of complementary expertise and efficient use of personnel, time, and resources. Progress has been made in this area, with SOPs, MAPPs, etc. based on mandates from FDAAA, as well as addressing concerns raised in the 2009 GAO report. Further clarification of accountability, establishment of "best practices", and evaluation of outcomes of newly mandated reviews (BPCA, FDAAA) are needed.

Similarly, there needs to be clarity about evolving technologies of data mining, Sentinel, etc. We found considerable variability in understanding of when, by whom, how, and why to do data mining, as well as uncertainty of the relative roles of Sentinel and spontaneous reports among staff. Leadership and education throughout will be needed to optimize FDA's public health roles in drug safety. In determining roles and accountabilities, it is obvious that staff and leadership from throughout FDA bring expertise and excellence, as well as potential biases based on their backgrounds, training, and current responsibilities with respect to a given medicine. This can be viewed as a positive for creative discourse leading to regulatory action, but "ownership" and primary responsibilities for various aspects of data collection, analysis and interpretation still need to be clarified, and converted into best practices, with recognition and respect of all approaches to assessing benefit:risk.